AD-A278 164

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1994	Professional Paper

4. TITLE AND SUBTITLE

A THEORY OF CONDITIONAL INFORMATION WITH APPLICATIONS

5. FUNDING NUMBERS

In-house funding

6. AUTHOR(S)

P. G. Calabrese

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Command, Control and Ocean Surveillance Center (NCCOSC)
RDT&E Division
San Diego, CA 92152–5001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Naval Command, Control and Ocean Surveillance Center (NCCOSC)
RDT&E Division
San Diego, CA 92152–5001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The development of conditional propositions, deduction between conditionals, and boolean-like operations on conditionals, and their associated probabilities are here unified in terms of boolean relations of the form "b = 0" defined on an initially relation-free algebra of boolean polynomials that transcend an initial domain of discourse. The conditional proposition (a/b) is assigned the conditional probability P(a/b), which is different from the probability that (a/b) is a tautology. The resulting algebraic techniques are demonstrated in several examples such as by simplifying a circular rule-based expert system and removing its circularity and by deriving logical and probability formulas for keeping communication lines open.

Published in *IEEE Transactions of Systems, Man, and Cybernetics*, December 1993, pp. 1–10.

14. SUBJECT TERMS

conditional propositions
conditionals

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAME AS REPORT

UNCLASSIFIED

21a. NAME OF RESPONSIBLE INDIVIDUAL	21b. TELEPHONE (include Area Code)	21c. OFFICE SYMBOL
P. G. Calabrese	(619) 553-4042	Code 421

NSN 7540-01-280-5500

Standard form 298 (BACK)

UNCLASSIFIED

For publication in IEEE Tr SMC

A Theory of Conditional Information with Applications

Philip G. Calabrese

Abstract

The development of conditional propositions, deduction between conditionals, and boolean-like operations on conditionals, and their associated probabilities are here unified in terms of boolean relations of the form "b = 0" defined on an initially relation-free algebra of boolean polynomials that transcends an initial domain of discourse. The conditional proposition (a|b) is assigned the conditional probability P(a|b), which is different from the probability that (a|b) is a tautology. The resulting algebraic techniques are demonstrated in several examples such as by simplifying a circular rule-based expert system and removing its circularity and by deriving logical and probability formulas for keeping communication lines open.

Introduction

Among the various concepts[1] residing at the nexus of so many intellectual subtleties that have come into scientific consciousness because of our efforts to replicate human information processing in computers (Artificial Intelligence), none is more daunting than the central notion of a conditional proposition (in logic), a conditional event (in probability) - conditional information in general.

Conditional Information

Actually, all information is inherently conditional: It is quite impossible to propound a proposition without assumptions! These assumptions are the conditions or context of the information. However, some assumptions are *always* implicit rather than explicit. This fact of intellectual life is enshrined by science and philosophy in the axiomatic method, which ever seeks to prove consequences *from* assumed postulates. The final premise of Science is material existence and knowledge thereof; the final premise of religion is God and the knowledge thereof. The premise of cosmology is the whole of material reality and the knowledge thereof; the premise of philosophy is the whole of *all* reality and the ability to know it. So it is impossible to escape from conditions. Conditional propositions are a basic intellectual element more so than so-called "unconditioned" boolean propositions.

Another way to say this is that all information has context, which is just another word for conditions or premises. So in combining information with differing contexts, those contexts must be carried along with the consequent action (or equivalent proposition) that together form the premise-conclusion pair. This is a fundamental unit of thought, every proposition is really such a pair. When a set of propositions has a common context we tend to drop the context and represent just the consequent proposition. Thus an infinite regress is avoided. In the final analysis all information has a common context, namely the whole universe of mind & matter in which that information fits and has its meaning.

Truth, Implication and Uncertainty. Now matters are well understood as long as conditional propositions are evaluated as being "true" according as they are true in *all* cases, but false if even one counterexample can be found, one case in which the antecedent is true but the consequent is false. A conditional "if b then a" is *wholly* true, true in all cases, a theorem, a so-called tautology, *necessarily* true - whenever every instance (case) of b is also an instance of a. In symbols, $b \leq a$. Letting $\wedge$ mean "and", $\vee$ mean "or" and $\neg$ mean "not", this can also be said in several other equivalent ways: "In all cases either proposition a is true or proposition b is false"; that is, $a \vee \neg b = 1$. Equivalently, $a \wedge b = b$, $a \vee b = a$, $\neg a \wedge b = 0$.

But when some uncertainty is introduced, the subtle difficulties that arise become almost insuperable. For instance, while standard probability theory has quite adequately quantified the partial truth in "unconditioned" statements residing in any measurable boolean algebra, the treatment of *conditionals* in standard probability theory is decidedly rudimentary.[2] There is only a "conditional probability", P(a|b), of "a given b", which is the ratio of

Manuscript received , 1992. This work was done while the author held a National Research Council Senior Research Associateship at the Research, Development, Test and Evaluation Division (NRaD), Code 421, Naval Command, Control and Ocean Surveillance Center (NCCOSC) in San Diego, California 92152. Financial support is gratefully acknowledged from Mr. Michael Mudurian of the Joint Directors of the Laboratories and Dr. Ralph Wachter of the Office of Naval Research.

[1] L. Zadeh [21] has explored the inherent "fuzziness" (imprecision) of many of our most useful words; G. Shafer [20] has explored the ubiquitous uncertainty in information in terms of "support" for "beliefs". The approach here is more classical and perhaps less ambitious, being an extension of boolean algebra and probability theory to conditionals.

[2] Those who have made serious attempts, with some success, to algebraically develop conditionals include G. Boole [2], B. Russell [18], B. De Finetti [8], Ernest Adams [1], G. Schay [19], D. Lewis [14], N. Nilsson [15], P. Calabrese [3], T. Hailperin [10] and I.R. Goodman, H.T. Nguyen & E. Walker [12] and others, but the list is not long.

the probabilities of the instances of $(a \wedge b)$ to the probability of the instances of b. While the conditional probability, P(a|b), is essential to probabilists for quantifying the partial truth of (a|b), its divergence[3] from $P(a \vee \neg b)$ whenever P(b) and P(a|b) are both less than 1 dramatizes the fact that new distinctions must be made in the realm of conditionals and new structures identified as soon as non-trivial amounts of uncertainty are incorporated. No longer can the material conditional proposition $(a \vee \neg b)$ be used to measure the truth of a conditional (a|b) in all circumstances.[4] It is also unfortunate that so many authors refer to $(a \vee \neg b)$ as "material implication" when in fact $(a \vee \neg b)$ is not an implication - notwithstanding that it is often expressed as "$b \supset a$"; rather, $(a \vee \neg b)$ is a proposition or an event. By contrast, material implication is a *relation* defined by the boolean equation $(a \vee \neg b = 1)$ or equivalently by $(\neg a \wedge b = 0)$. There is all too little explicit distinction made between absolutely true statements and higher order statements of the absolute truth of a statement. The statement that proposition b is wholly true is not the same as the (wholly true) statement b. This distinction is crucial when it comes time to put a separate probability measure on the various possible deductively closed sets of propositions and conditional propositions that may presently form the assumed context of some other set of uncertain propositions & conditional propositions. This is appropriate in 2-stage experiments or in other time-indexed information processing.

An Algebra of Conditional Information. Unlike the standard probability theory of unconditioned propositions, there is no standard algebra of conditionals by which to manipulate or simplify a complex expression involving conditionals before performing a probability calculation. Some researchers believe that it is unnecessary to combine conditionals directly by "and" and "or". Others see no need for iterated (nested) conditionals.

Yet, even so common a parental admonition as "In the rain wear a raincoat, and when it's cold wear warm clothes" is already a conjunction of two conditionals which can be violated in either of two ways and satisfied by wearing an insulated raincoat in cold rainy weather. When relationships exist between the components of conditionals, it can hardly be right to ignore them in our logical and probabilistic calculations. Furthermore, nested conditionals are essential for deduction between conditionals and for conditioning upon new possibly conditional information. Finally, we all routinely use such constructions as "if whenever you walk across the street you look both ways, then you will cross safely."

In this regard E. Adams [1] has posed an interesting example: An object of unknown color may be red (r), yellow (y) or blue (b) with equal probability. What is the new probability of blue upon learning that 'if the object is not red it is blue'? Thus, $P(\, b \mid (b \mid \neg r)\,) = ?$ and more importantly, how should this probability be calculated?

Conditionals also arise naturally in expert systems, which tend to be organized in terms of "if - then -" rules. Difficulties arise when these sets of rules are circular or inconsistent; simplification can identify and eliminate redundancies and inconsistencies and can also relieve computational complexities. Conditionals can also be useful in managing data bases, combining data, facilitating queries and quantifying partial truth.

Overview. A section on finite theory, illustrated with simple examples, is followed by an applications section containing somewhat more elaborate examples. The theory section provides a new, fundamental approach to conditional propositions as residue classes of boolean relations of the form (b = 0) on an initially *relation-free* boolean polynomial domain as generated by a finite set of propositions of interest. This allows a new, unified development of conditional propositions and iterated conditionals and even deduction between conditionals, as well as boolean-like operations thereon, as previously formulated in Calabrese [4,5,6]. Conditionals that are themselves conditions for some other proposition act like their corresponding material conditionals. The concepts of conditional implication ($\leq_c$) and probabilistically-monotonic implication ($\leq_m$) also reduce to boolean relations of this same form and thus are available for building deductively closed systems (algebraic filters) of conditional propositions. The difference between P(c) versus P(c = 1), and between $P(c \vee \neg d)$ versus $P(c \vee \neg d = 1)$ is algebraically defined in preparation for a two-stage experiment in which these must be distinguished. Conditional propositions are combined by "and", "or" and "not" in a fresh account based on boolean filter theory and boolean relations. The probability formula $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$ is generalized to conditionals in preparation for its later use in a communication link example. The applications section includes three expert system (if - then -) rules that result in a circular, rule-chaining system but which upon being combined into a single rule, need no further chaining, thereby avoiding an infinite do-loop at execution time.

Finite Theory of Conditional Information

The development of conditional propositions presented here will incidentally avoid the triviality situations exhibited by D. Lewis [14] by embedding the original boolean propositions in a larger, generally *non-boolean,*

---

[3] This difference is $P(a \vee \neg b) - P(a|b) = (1 - P(b))(1 - P(a|b))$ as expressed by P. Calabrese [3] in 1974.

[4] D. Lewis [14] early showed that (a|b) could not be assigned the probability P(a|b) and also be an element of the original boolean algebra containing a and b because then $P(a|b) = P((a|b) \wedge a)) + P((a|b) \wedge \neg a)) = P((a|b) \mid a)\, P(a) + P((a|b) \mid \neg a) = \ldots = 1 \times P(a) + 0 \times P(\neg a) = P(a)$, no matter what (except for trivial cases) the propositions a & b.

space of ordered pairs of propositions. As exhibited in P. Calabrese [3,4,5,6,7], the structure of these ordered pairs (conditional propositions) allows extended operations of conjunction ($\wedge$), disjunction ($\vee$), negation($\neg$) and conditioning (|) in a way that is non-trivially consistent with assigning each conditional proposition or event, (a|b) the probability P(a|b); the probability that (a|b) is a tautology may have some other probability.

Besides the ordered-pairs construction, conditional propositions can be algebraically defined in several other equivalent ways and there is also more than one plausible (non-equivalent) way to define operations on the resulting conditionals. For an overview of these approaches see D. Dubois & H. Prade [9] and I. R. Goodman et al [12,13]. The approach here is consistent in spirit with earlier work of G. Schay [19] and E. Adams [1], but also differs extensively in content and emphasis. This is still a young topic.

Propositions and Events

For ease of application, especially in computers, we restrict attention to the finite case. Let S be a finite set of n propositions $a_1, a_2, ..., a_n$ of interest. Assume that none of the negations, $\neg a_i$, of these a_i appears elsewhere in the list of n initial propositions. In order to transcend the domain of discourse S, form $\mathbf{B}(x_1, x_2, ..., x_n)$, the *free* boolean algebra of polynomials generated by these n propositions (when taken as relation-free variables) under negation ($\neg$), conjunction ($\wedge$), and disjunction ($\vee$). Thus $\mathbf{B}$ is the set of all finite disjunctions of conjunctions of n free propositional variables $x_1, x_2, ..., x_n$ or their negaions. An arbitrary element of $\mathbf{B}$ is the finite disjunction of any subset of (non-zero) *atomic propositions* ω of the form $e_1x_1 \wedge e_2x_2 \wedge ... \wedge e_nx_n$ where, for $i \le n$, each e_ix_i is either x_i or $\neg x_i$. Let Ω be the set of all such atoms ω. Ω has 2^n elements and so $\mathbf{B}$, which corresponds to the collection of all subsets of atoms of Ω, contains 2 raised to the power 2^n propositions. The propositions of $\mathbf{B}$ also correspond to the collection of all possible ordered 2^n-tuples of 0s and 1s. Alternately, $\mathbf{B}$ can be expressed as all possible two-valued functions from Ω to $\{0,1\}$. In the standard way assign probabilities to the atoms of Ω and so to the propositions of $\mathbf{B}$.

For example, suppose S = {Blue (b), Red (r), Yellow (y)} represents the possible colors of an object. Then $\Omega = \{b \wedge r \wedge y, \neg b \wedge r \wedge y, b \wedge \neg r \wedge y, \neg b \wedge \neg r \wedge y, b \wedge r \wedge \neg y, \neg b \wedge r \wedge \neg y, b \wedge \neg r \wedge \neg y, \neg b \wedge \neg r \wedge \neg y\}$, which will be abbreviated by dropping the $\wedge$ symbols and writing b' for $\neg$b. So $\mathbf{B}$ contains all of the 2^8 subsets of Ω. Initially assign probability 1/8 to each of the elements of Ω.

With respect to machine storage and computation time, an n-tuple of binary storage bits requires only n ordered binary storage places of computer memory and yet is capable of taking the value of any one of 2^n atoms. Unfortunately, a proposition requires an ordered 2^n-tuple to be specified, corresponding to that subset of the 2^n atoms which the proposition contains. For n = 10, that requires a binary vector of length 1024. Thus for purposes of computational tractability it may be necessary to divide large numbers of propositions into groups of related propositions with no group having more than a specified maximum. Alternately, or in addition, a judicious choice of hierarchic variables (objects) can be defined to manage computational complexity.

Conditional Propositions

The conditional propositions (ordered pairs of propositions) will now be generated. In the context set forth above, the process of conditioning a proposition c upon another proposition b can be performed by simply setting to 0 all atoms of c in common with $\neg$b.

Conditionalization Relations R_b. More formally, temporarily fix an arbitrary proposition b in $\mathbf{B}$ and form the quotient boolean algebra ($\mathbf{B}$/b) under the congruence relation R_b defined by the boolean equation ($\neg b = 0$), as follows: For any two propositions c and d in $\mathbf{B}$, the ordered pair (c,d) is defined to be in the relation R_b if c and d are equal after setting to 0 any atoms that they have in common with $\neg$b. This amounts to specifying $\neg$b to be impossible, which is also to say that b is to be necessary. When propositions are represented as two-valued functions from Ω to $\{0,1\}$, this is equivalent to setting $b^{-1}(0) = \{\omega \in \Omega: b(\omega) = 0\} = \{0\}$, which is a form that easily extends to conditioning by conditionals.

Clearly, R_b is an equivalence relation and a congruence relation (=) using the fact that two propositions are equal if and only if they have the same atoms.

Denote by ($\mathbf{B}$/b) the quotient boolean algebra thus formed, and denote by (c|b) the residue class in ($\mathbf{B}$/b) containing the proposition c.[5] It follows that (c|b) = (d|b) if and only if $c \wedge b = d \wedge b$. We interpret (c|b) as "c given b" or "c in case b is true" or "c given $\neg$b is false" or as "c modulo ($\neg b = 0$)", or "c modulo (b = 1)". Alternately, (c|b) is c "given the falsity of $\neg$b" or "c given the truth of b".

Note that with respect to the set of all atoms Ω, (c|b) is a two-valued function *restricted* to the atoms in b. (c|b)

[5] If b is already 0 before forming ($\mathbf{B}$/b) then ($\mathbf{B}$/b) degenerates to the singleton {0}, which is the "inconsistent boolean algebra" in which 1 = 0.

assigns 1 to those atoms in $(c \wedge b)$ and 0 to the atoms in $(\neg c \wedge b)$. (c|b) is undefined (or "inapplicable") on the atoms in $\neg b$ because the construction of (c|b) assumes that they are no longer atoms having been set to 0, which is not an atom. So, as early pointed out by B. De Finetti [8], with respect to the original boolean algebra $\mathbf{B}$, a conditional proposition (c|b) has three truth "states": 1, 0 and "undefined". For any proposition c, the probability of (c|b) is set to $P(c \wedge b) / P(b)$ so that $P(b|b) = 1$.

The congruence class containing c is $(c|b) = \{x \in \mathbf{B}: x \wedge b = c \wedge b\} = \{x \in \mathbf{B}: x = (c \wedge b) \vee (\neg b \wedge y)$ for some y in $\mathbf{B}\} = cb \vee \neg b\mathbf{B}$, where conjunction has been replaced by juxtaposition and $\neg b\mathbf{B} = \{\neg b \wedge y: y \in \mathbf{B}\}$.

The congruence class $(b|b) = \{x \in \mathbf{B}: x \wedge b = b \wedge b\} = \{x \in \mathbf{B}: b \leq x\}$ is a boolean filter, a deductively closed set of propositions, because the class is closed under conjunction, and also closed under boolean deduction $\leq$.

The quotient boolean algebra $\mathbf{B}/b$ inherits a natural deduction ordering $\leq$ from $\mathbf{B}$, namely that $(a|b) \leq (c|b)$ if and only if $ab \leq cb$.

The Conditional Closure of $\mathbf{B}$. Let $\mathbf{B}/\mathbf{B}$ denote the set of all conditional propositions (a|b) for arbitrary propositions a and b in $\mathbf{B}$. $\mathbf{B}/\mathbf{B}$ will be called the conditional closure of $\mathbf{B}$. Note that $\mathbf{B}$ is isomorphic to $\mathbf{B}/1$, so that $\mathbf{B}/\mathbf{B}$ includes a representative of $\mathbf{B}$. Two conditional propositions (a|b) and (c|d) are equivalent in $\mathbf{B}/\mathbf{B}$ if $(b = d)$ and $(a \wedge b = c \wedge d)$. Thus in general $(a|b) = (a \wedge b \mid b)$, which has been called "reduction". However, the further extension of this equation to cond-itionals as $(a|b) \mid (c|d) = [(a|b) \wedge (c|d)] \mid (c|d)$ is questionable since $(a|b) \mid (c|d)$ can also be rendered as $(a \mid b \wedge (c|d))$.

Continuing the colored object example, if it is known that the object has exactly one of the 3 colors then the free boolean algebra $\mathbf{B}$ must be conditioned by the information that $(br \vee by \vee ry \vee b'r'y') = 0$. So Ω becomes $\{br'y', b'yr', b'y'r\}$ and $\mathbf{B}$ now has $2^3 = 8$ non-zero elements. In that situation the (conditional) probability of each of the three non-zero elements becomes $(1/8) / (3/8) = 1/3$.

Iterated (Nested) Conditionals. The virtue of this manner of development of $\mathbf{B}/b$ via boolean relations defined by a boolean equation $(\neg b = 0)$ is that the conditionalization process can be naturally extended to iterated conditional propositions of the general form $(a|b) \mid (c|d)$. While there may be other ways to perform iterated conditionalization, the intention here is to explore the following type: Once $\mathbf{B}(x_1, x_2, \ldots, x_n)$ has been constructed to transcend the initial set S of n propositions, all subsequent conditioning is applied to this original polynomial boolean algebra $\mathbf{B}$. Thus all subsequent conditions are applied to $\mathbf{B}$, and hence to all structures built from $\mathbf{B}$ including $\mathbf{B}/\mathbf{B}$. Of course if new propositions are added to the initial set of propositions S, then $\mathbf{B}$ must be expanded accordingly.

Consider first the special case $((a|b) \mid c)$. (a|b) is the proposition a with all of its atoms in common with $\neg b$ set to 0, and so $(a|b) \mid c$ is the resulting proposition after all of its atoms in common with $\neg c$ are also set to 0. But $(\neg b = 0)$ and $(\neg c = 0)$ if and only if $(\neg b \vee \neg c = 0)$. So $((a|b) \mid c)$ is the proposition a with all atoms in common with $(\neg b \vee \neg c) = \neg(b \wedge c)$ set to 0. So $((a|b) \mid c) = (a \mid (bc))$. This has been called the "import-export" principle.

Iterated Conditionals of the form $((a|b) \mid c)$. More formally, for any two conditional propositions (a|b) and (g|h) in $\mathbf{B}/\mathbf{B}$, the ordered pair ((a|b), (g|h)) is defined to be in the relation R_c if (a|b) and (g|h) are equal after setting to 0 any atoms that their components a, b, g and h have in common with $\neg c$. Clearly this forms an equivalence relation on $\mathbf{B}/\mathbf{B}$. It easily follows that modulo R_c each conditional proposition (a|b) is equivalent to (ac|bc), which is equivalent to (a|bc). Thus $((a|b) \mid c) = ((a \mid bc) \mid c)$, which is just $(a \mid bc)$ because $\neg c \leq \neg(bc)$ and so the atoms of $\neg c$ are already zeroed as a part of the atoms of $\neg(bc)$ being zeroed. It also follows that R_c is a congruence relation on $\mathbf{B}/\mathbf{B}$ because if $(a|b) = (g|h)$, and so $b = h$ and $ab = gh$, then $bc = hc$ and $abc = ghc$, and so $(a|bc) = (g|hc)$.

Iterated conditionalization by successive propositions amounts to setting additional atoms in Ω to 0 and thus the process stays inside $\mathbf{B}/\mathbf{B}$, which contains all possible such zeroing of atoms.

Another way to formally derive that $((a|b) \mid c) = (a \mid bc)$ is via the algebraic congruence class representation: $((a|b) \mid c) = ((ab \vee \neg b\mathbf{B}) \mid c) = (ab \vee \neg b\mathbf{B})c \vee \neg c\mathbf{B} = abc \vee \neg bc\mathbf{B} \vee \neg c\mathbf{B} = abc \vee (\neg bc \vee \neg c)\mathbf{B} = abc \vee (\neg b \vee \neg c)\mathbf{B} = abc \vee \neg(bc)\mathbf{B} = (abc \mid bc) = (a|bc)$.

Iterated Conditionals of the form $(a \mid (c|d))$. To say "given (c|d)" or "in case (c|d)" is to assume the non-violation of (c|d). Since (c|d) is violated on $(\neg c \wedge d)$, this amounts to setting to zero all atoms of proposition a in common with $(\neg c \wedge d)$. So $(a \mid (c|d)) = (a \mid (c \vee \neg d))$ because conditioning by $(c \vee \neg d)$ has the same result.

Conditionalization Relations $R_{(c|d)}$. More formally, for any two propositions a and b in $\mathbf{B}$, the ordered pair (a,b) is defined to be in the relation $R_{(c|d)}$ if a and b are equal after setting to 0 any atoms that they have in common with $(c|d)^{-1}(0)$, which is the set of atoms for which (c|d) is violated. As in the simple conditional case, this means that we must set $(c|d)^{-1}(0) = \{0\}$. However, $(c|d)^{-1}(0) = \{\omega \in \Omega: \omega \text{ is an atom of } (\neg c \wedge d)\}$. So the atoms of $(\neg c \wedge d)$ are just the atoms to be set to zero, and thus $(a \mid (c|d)) = (a \mid (c \vee \neg d))$ since $(c \vee \neg d) = \neg(\neg c \wedge d)$. It is easy to see that $R_{(c|d)}$ is a congruence relation on $\mathbf{B}$.

For example, returning to E. Adam's colored object example, upon learning that "if the object is not red then it is blue", the new (conditional) event for "blue (b)" is $(b \mid (b \mid \neg r)) = (b \mid (b \vee r))$, and the new probability for b is $P(b \mid (b \vee r)) = P(b) / P(b \vee r) = (1/3) / (2/3) = 1/2$.

Iterated Conditionals of the form ((a|b) | (c|d). On the face of it, by observing the arrangement of parentheses, ((a|b) | (c|d) is a conditioning of proposition a by proposition b followed by a conditioning of the result by (c|d). Thus ((a|b) | (c|d) is proposition a after its atoms in common with $\neg b$ are set to zero and then after its atoms in common with $(\neg c \wedge d)$ are also set to zero, which is just $[(a \mid b) \mid (c \vee \neg d)] = (a \mid (b \wedge (c \vee \neg d)))$, and the latter conditional proposition is in $\mathbf{B}/\mathbf{B}$. Thus $\mathbf{B}/\mathbf{B}$ is closed under such iterated conditioning. We have then that

$$((a|b) \mid (c|d) = (a \mid (b \wedge (c \vee \neg d))) \qquad (1)$$

Boolean Relations. The conditional closure $\mathbf{B}/\mathbf{B}$ can be conditioned by any boolean relation which is expressible in the form "b = 0". These allow known relationships between the propositions of $\mathbf{B}$ such as "$c \le d$", "$cd = 0$", and "$c \vee d = 1$" and similar relations on conditionals to be incorporated in terms of conditionalization.

This also opens up connections between this theory of conditional event algebra / conditional probability (CEAPL) and the work of J. Hooker [11] concerning techniques for solving systems of boolean equations; in many situations, a set of these boolean equations will need to be simultaneously solved in order to "define the condition" as specified by a set of known boolean relationships between specified propositions and conditional propositions.

Deduction

Since a conditional proposition (c|d), *when acting itself as a conditional*, is equivalent to its corresponding material conditional $(c \vee \neg d)$, a collection $(a_1|b_1)$, $(a_2|b_2)$, $(a_3|b_3)$, ..., $(a_m|b_m)$ of conditionals generates an algebraic filter, namely the deductively closed system generated by the proposition $(a_1 \vee \neg b_1) \wedge (a_2 \vee \neg b_2) \wedge (a_3 \vee \neg b_3) \wedge \ldots \wedge (a_m \vee \neg b_m)$. In particular, given one conditional (a|b), a second conditional (c|d) is also deductively "given" (that is, a tautology) if and only if $(a \vee \neg b) \le (c \vee \neg d)$. This kind of implication between conditionals has been called *conditional implication* ($\le_c$). As defined in [6],

$$(a|b) \le_c (c|d) \text{ if and only if } (a \vee \neg b) \le (c \vee \neg d). \qquad (2)$$

$(a|b) \le_c (c|d)$ means that "if (a|b) is not false then (c|d) is not false." Equivalently, it means that "if (c|d) is false then (a|b) is false."

Note that with respect to probability, conditional implication ($\le_c$) only ensures that $P(\neg cd) \le P(\neg ab)$ or equivalently that $P(a \vee \neg b) \le P(c \vee \neg d)$ not that $P(a|b) \le P(c|d)$, which is stronger, in general requiring also that $ab \le cd$. When both $(a|b) \le_c (c|d)$ and $ab \le cd$ hold, the implication has been called *monotonic* implication by the author [6] since it is probabilistically monotonic.[6] The conditional implication $(a|b) \le_c (c|d)$ can be expressed by the boolean equation $(\neg cd)(a \vee \neg b) = 0$ and monotonic implication $(a|b) \le_m (c|d)$ can be expressed, for instance, by the equation $(\neg cd)(a \vee \neg b) \vee (ab)\neg(cd) = 0$.

For Adams [1], (a|b) can be "validly inferred" from (c|d) if and only if (c|d) has high conditional probability whenever (a|b) has high conditional probability, and this is equivalent to monotonic implication, but Adams generalizes by allowing the premise to be any finite conjunction of conditionals.

Two conditionals (a|b) and (c|d) are said to be *conditionally equivalent* ($=_c$), that is, equivalent as conditions, if and only if $(a \vee \neg b) = (c \vee \neg d)$.

If $(a|b) \le_c (c|d)$ then (c|d) | (a|b) is a tautology because $(c|d) \mid (a|b) = (c \mid d \wedge (a \vee \neg b)) = (c \mid d \wedge (c \vee \neg d)) = (c \mid d \wedge c) = (d \wedge c \mid d \wedge c) = (1 \mid d \wedge c)$. Conversely, if (c|d) | (a|b) is a tautology then $(a \vee \neg b) \le (c \vee \neg d)$ because then $(c|d) \mid (a|b) = (1|d) \mid (a|b)$ and so $cd(a \vee \neg b) = d(a \vee \neg b)$, which means $d(a \vee \neg b) \le c$, which means $(a \vee \neg b) \le (c \vee \neg d)$.

Note that since $(c \vee \neg d) = \neg d \vee \neg(\neg c)$, a conditional proposition (c|d), "if d then c", is conditionally equivalent to its contrapositive $(\neg d \mid \neg c)$, "if not c then not d". $(c|d) =_c (\neg d \mid \neg c)$ because $(d \le c)$ if and only if $(\neg c \le \neg d)$. This is, after all, a fact of everyday mathematical life (except for those who resolutely abhor all proofs by contradiction.) However, $(c|d) \ne (\neg d \mid \neg c)$ and $P(c|d) \ne P(\neg d \mid \neg c)$ except in very special cases.

Since by definition every deductively closed set $\mathbf{D}$ contains $(a|b) \wedge (c|d)$ whenever it contains both (a|b) and (c|d), a set

[6] I.R. Goodman [12] first showed that this implication relation between conditionals is also the weakest which is probabilistically monotonic for all probability measures.

of m conditionals deductively entails all 2^m of its subset conjunctions. $\mathfrak{D}$ also contains (c|d) whenever it contains (a|b) and (a|b) $\leq_x$ (c|d) holds, where relation $\leq_x$ is $\leq_c$, $\leq_m$, or some other boolean deduction relation in the sense that it extends $\leq$ on $\mathfrak{B}$/b for every b $\in$ $\mathfrak{B}$. That is, (a|b) $\leq_x$ (c|b) if and only if ab $\leq$ cb. Furthermore, modus ponens is also valid in the conditional closure in the sense that (a|b) $\wedge$ [(c|d) | (a|b)] = (a|b) $\wedge$ (c|d).

All this reduces the problem of deduction in the realm of conditionals to the well-understood boolean deduction, much like the system of N. Nilsson [15], in which *all* conditionals (a|b), not just those acting as conditions, are reduced to their corresponding material conditionals (a $\vee$ $\neg$b). In this context, probabilities of conditionals (c|d) are probabilities of their corresponding material conditionals (a $\vee$ $\neg$b) since both are false on the same set of atoms, namely those of ($\neg$a $\wedge$ b). That is, P((a|b)) = P(a $\vee$ $\neg$b).

Tautologies. In view of the above, a probability can be either of the following two important types:

1) Probability that a proposition is true, and
2) Probability that a proposition is a tautology.

For a proposition c, these types of probability are

1) P(c) = P{$\omega \in \Omega$: *ω is an atom of proposition c*}
= P(c^{-1}(1)), and
2) P((c)) = P(proposition c is a tautology) = P(c = 1)
= P(principle filter generated by (c) is given).

For conditionals, the first type is P(c|d) = P(cd) / P(d) = P($(cd)^{-1}$(1)) / P(d^{-1}(1)), called the conditional probability, while the second type is P((c|d) is a tautology) = P(d $\leq$ c) = P(c $\vee$ $\neg$d = 1).

Again, in the second type, P((c|d) = (1|d)) = P((c $\vee$ $\neg$d) = 1), as in the standard logical reduction, because (c|d) is a tautology if and only if (c $\vee$ $\neg$d) is a tautology. In the first type, P(c|d) < P(c $\vee$ $\neg$d) except when P(d) or P(c|d) is 1. Not noticing this distinction has been a great source of confusion between logicians and probabilists.

This requires P to be extended to the space of filters of $\mathfrak{B}$, which form a boolean algebra. For filters J and K,

$$\begin{aligned} (J \vee K) &= \{a \vee b: a \in J, b \in K\} = J \cap K \\ (J \wedge K) &= \{a \wedge b: a \in J, b \in K\} \supseteq J \cup K \\ J' &= \{b \in \mathfrak{B}: \neg a \leq b, \text{ for all } a \in J\} \\ (0) &= \mathfrak{B} \\ (1) &= \{1\} \end{aligned}$$

Then J $\wedge$ J' = (0), J $\vee$ J' = (1), and (J')' = J. This allows P to apply also to the likelihood of tautologies, which is different from the likelihood of propositions. Deduction between filters "if J is a tautology then K is a tautology" has been called necessary implication ($\leq_n$), the "entailment of necessity": if J = (1) that K = (1). In general (a|b) $\leq_n$ (c|d) means that (b $\leq$ a) entails (d $\leq$ c).

Other Operations on Conditionals

As mentioned previously, operations of negation ($\neg$), conjunction ($\wedge$) and disjunction ($\vee$) have been plausibly defined on the conditionals of $\mathfrak{B}/\mathfrak{B}$ in more than one way. I. R. Goodman [12] has shown that choosing such an operation on $\mathfrak{B}/\mathfrak{B}$ corresponds to choosing a 3-valued truth table on $\mathfrak{B}/\mathfrak{B}$ and conversely. Here, the main consideration is to be consistent with probability while developing an algebra of logical conditionals which is consistent with common usage. The spirit follows the development of conditionals by E. Adams [1] although it also differs significantly. For instance, recently Adams has focused on meta-conditionals like P(P(c|d) $\approx$ 1)), the probability that c is highly likely given d. This is a generalization of P((c|d) = 1), the probability that (c|d) is a tautology.

Negation. Within any quotient boolean algebra $\mathfrak{B}$/b, each conditional proposition (a|b) has a negation $\neg$(a|b), which is just ($\neg$a|b) because (a|b) $\vee$ ($\neg$a|b) = (1|b), and (a|b) $\wedge$ ($\neg$a|b) = (0|b), where (1|b) and (0|b) are the 1 and 0 respectively of $\mathfrak{B}$/b. Furthermore, P(a|b) + P($\neg$a|b) = 1. So it is natural to keep these conditional negations in $\mathfrak{B}$/b:

$$\neg(a|b) = (\neg a|b) \tag{3}$$

However (a|b) and ($\neg$a|b) are not negations with respect to each other in $\mathfrak{B}/\mathfrak{B}$ except in the case that b = 1. The negation of a (principal) filter generated by the single proposition b is ($\neg$b) = {x $\in$ $\mathfrak{B}$: $\neg$b $\leq$ x} and the congruence class containing $\neg$a is ($\neg$a|b).

Disjunction. Intuitively speaking, the compound conditional "if b then a, or if d then c", in symbols (a|b) $\vee$ (c|d), is true in case either conditional is true, and it is defined (or "applicable") in case either antecedent, b or d, is true. If (a|b) $\vee$ (c|d) is defined but not true then it is false. Since (a|b) is true on ab and (c|d) is true on cd, it follows that

$$(a|b) \vee (c|d) = (ab \vee cd \mid b \vee d). \tag{4}$$

More formally, the deductively closed set (or filter) associated with the conditional proposition (a|b) is (b) = {x $\in$ $\mathfrak{B}$: b $\leq$ x} and that of (c|d) is (d) = {x $\in$ $\mathfrak{B}$: d $\leq$ x}. If *either* of these two deductive systems is "given", then the resulting deductive system that is "given" necessarily includes only those propositions that are common to (b) and (d). But the largest filter included in both filters (b) and (d) is $(b) \cap (d) = (b) \vee (d) = (b \vee d)$. So $(a|b) \vee (c|d)$ is defined on $(b \vee d)$. Within this domain of definition, $(a|b) \vee (c|d)$ is true on $(ab \vee cd)$. So equation (4) follows. Alternately, if all atoms in common with $\neg b$ are set to zero, or all atoms in common with $\neg d$ are set to zero, then the atoms that are actually set to zero in either case are just those in common with $(\neg b \wedge \neg d)$. But $(\neg b \wedge \neg d) = 0$ if and only if $\neg(b \vee d) = 0$. So again, $(a|b) \vee (c|d)$ is defined on $(b \vee d)$.

Conjunction. As in the case of disjunctions, $(a|b) \wedge (c|d)$ is defined when either $(a|b)$ or $(c|d)$ is defined, that is, on $(b \vee d)$. Within its domain of definition, $(a|b) \wedge (c|d)$ is violated if and only if either $(\neg a \wedge b) \vee (\neg c \wedge d)$ is true. Otherwise, on $\neg[(\neg a \wedge b) \vee (\neg c \wedge d)] = (a \vee \neg b)(c \vee \neg d)$, it is not violated. If it is defined and not violated, then it is true. It follows that

$$\begin{aligned}(a|b) \wedge (c|d) &= (a \vee \neg b)(c \vee \neg d) \mid (b \vee d) \qquad (5)\\ &= (abcd \vee ab\neg d \vee \neg bcd) \mid (b \vee d)\\ &= [(ab)(c \vee \neg d) \vee (a \vee \neg b)(cd)] \mid (b \vee d).\end{aligned}$$

Note that the right hand side of equation (5) reduces to that of equation (4) when $ab \le (c \vee \neg d)$ and $(cd) \le (a \vee \neg b)$. These are equivalent to saying $(ab)(\neg cd) = (\neg ab)(cd) = 0$. A sufficient condition for this is that $bd = 0$, that b and d are disjoint. So in these situations the conjunction of conditionals may be equivalent to the disjunction of conditionals without the component conditionals being equal.

The conjunction of two conditionals can also be derived by requiring that the De Morgan formulas should hold in $\mathbf{B}/\mathbf{B}$: In that case $(a|b) \wedge (c|d) = \neg[(\neg a|b) \vee (\neg c|d)] = \neg[(\neg ab \vee \neg cd) \mid (b \vee d)] = (a \vee \neg b)(c \vee \neg d) \mid (b \vee d)$, which reduces to $(abcd \vee ab\neg d \vee \neg bcd) \mid (b \vee d)$. Similarly, equation (4) can be derived from equation (5) and the De Morgan formula. Alternately, the De Morgan formulas can be proved from equations (4) and (5).

The disjunction and conjunction of conditionals constitutes a combining of partially applicable information, as when data bases are built up from overlapping pieces of information from different sources or circumstances.

For example, consider the familiar experiment of rolling a single die once with atoms $\Omega = \{1, 2, 3, 4, 5, 6\}$. $\mathbf{B}$ is the collection of subsets of Ω. A wager is made that "if the roll is ≤ 3 then it will be odd and if the roll is even then it will a 6." This can be represented as $(\text{odd} \mid \le 3) \wedge (6 \mid \text{even}) = (0 \vee (\text{odd})(\le 3)(\neg(\text{even})) \vee \neg(\le 3)(6)(\text{even})) \mid (\le 3 \vee \text{even}) = (\{1, 3\} \vee \{6\} \mid \{1, 2, 3, 4, 6\}) = \{1, 3, 6\} \mid \{1, 2, 3, 4, 6\}$. The conditional probability is therefore 3/5, which can also be determined by checking individual atoms.

According to these operations, the conjunction of two conditionals $(a|b)$ and $(c|d)$ conditionally implies either of the conjuncts. That is, $(a|b) \wedge (c|d) \le_c (a|b)$. However, the disjunction of two conditionals is not in general conditionally implied by each of its disjuncts. That is, $(a|b) \le_c (a|b) \vee (c|d)$ does not hold in general, requiring in addition that $\neg cd \le b$, namely that if $(c|d)$ is false then $(a|b)$ is defined. Nevertheless, $(a|b) \wedge (c|d) \le (a|b) \vee (c|d)$ since both are in $\mathbf{B}/(b \vee d)$. In this regard, it turns out that except for special cases, we generally have[7]

$$(a|b) \wedge (c|d) \;<\; (a|b) \vee [(a|b) \wedge (c|d)] \;<\; a|b) \vee (c|d)$$

where the middle expression is also in $\mathbf{B}/(b \vee d)$. This middle expression is true whenever $(a|b)$ is true, even if $(c|d)$ is false, or when $(a|b)$ is undefined and $(c|d)$ is true, but it is violated when $(a|b)$ is false and $(c|d)$ is true. Similarly for $(a|b) \wedge [(a|b) \vee (c|d)]$.

Note that this property, for example, allows $(0|b) \wedge 1 = b$ to have larger probability in general than $(0|b)$, which must have probability 0. Conjunction and disjunction of conditionals does not here preserve probability in the same way. $(0|b)$ is zero only in $\mathbf{B}/b$. *Outside b, it has no* affect; it is "inapplicable". Some of the results of the preceding pages can be summarized as follows:

Theorem: Let S be any finite set $\{a_1, a_2, ..., a_n\}$ of n boolean propositions, no two of which are negations of one another, and let P be a probability assignment to the 2^n (atomic) elements of Ω = {all conjunctions of the n members of S or their negations}. Then the relation-free boolean algebra $\mathbf{B}$ generated by n free boolean variables $x_1, x_2, ..., x_n$ has a conditional closure $\mathbf{B}/\mathbf{B}$ that includes an isomorphic copy of the boolean algebra generated by S; furthermore, $\mathbf{B}/\mathbf{B}$ is closed under the operations of conjunction ($\wedge$), disjunction ($\vee$), negation ($\neg$) and conditionalization (|) according to equations (3), (4), (5) and (1) respectively, and so any such operation on S is represented in $\mathbf{B}/\mathbf{B}$; and finally, for all $(a|b)$ in $\mathbf{B}/\mathbf{B}$, $P(a|b) = P(a \wedge b) / P(b)$.

Note that a formal disjunction of the congruence classes $(ab \vee \neg b\mathbf{B})$ and $(cd \vee \neg d\mathbf{B})$ does not result in equation (4), but instead yields $(ab \vee \neg b\mathbf{B}) \vee (cd \vee \neg d\mathbf{B}) = (ab \vee cd \vee \neg b\mathbf{B} \vee \neg d\mathbf{B}) = (ab \vee cd) \vee (\neg b \vee \neg d)\mathbf{B} = (ab \vee cd) \vee \neg(bd)\mathbf{B} = (ab \vee cd \mid bd)$, which is the disjunction

[7] $(a|b) <_x (c|d)$ means that $(a|b) \le_x (c|d)$ and $(a|b) \ne_x (c|d)$.

operation chosen by G. Schay [19] to go with the conjunction operation (5). This derivation separates the consequent (ab) from its antecedent $\neg b\mathbf{B}$ in a way that leads to the conditional $(ab \vee cd \mid bd)$, which is defined only when both antecedents are true. By similar development the formal conjunction of the two congruence classes $(ab \vee \neg b\mathbf{B})$ and $(cd \vee \neg d\mathbf{B})$ yields the conditional proposition $(abcd \mid bd \vee \neg ab \vee \neg cd)$, which is the conjunction operation favored by I.R. Goodman, et al [13]. This conditional, which is undefined outside $(b \vee d)$, and within $(b \vee d)$ is undefined on $(ab\neg d \vee \neg bcd)$ is the greatest lower bound of $(a|b)$ and $(c|d)$ with respect to $\leq_m$, and is also conditionally equivalent ($=_c$) to the right hand side of equation (5). The result of the disjunction operation of Goodman et al upon $(a|b)$ and $(c|d)$ is the conditional $(ab \vee cd \mid ab \vee cd \vee bd)$, which is the least upper bound with respect to $\leq_m$, and is also conditionally equivalent to Schay's disjunction mentioned above. When b and d are disjoint this disjunction yields a tautology, the filter generated by $ab \vee cd$, namely, $(1 \mid ab \vee cd)$. For instance, for the compound "if heads comes up then I win or if tails comes up then I lose" this disjunction yields the filter (and congruence class) of all propositions entailed by the proposition "heads comes up and I win or tails comes up and I lose". By contrast equation (4) yields the simple proposition "heads comes up and I win or tails comes up and I lose".

Probabilities. By taking probabilities of equations (4) and (5) for disjunction and conjunction and doing some rearranging, the following formulas for the disjunction and conjunction of conditionals are obtained in terms of standard conditional probabilities:

$$P((a|b) \vee (c|d)) = P(b \mid b \vee d)\, P(a|b) + P(d \mid b \vee d)\, P(c|d) - P(abcd \mid b \vee d) \tag{6}$$

$$P((a|b) \wedge (c|d)) = P(b \mid b \vee d)\, P(a\neg d|b) + P(d \mid b \vee d)\, P(c\neg b|d) + P(abcd \mid b \vee d) \tag{7}$$

Note that the last term in these two formulas can be written as $P(bd \mid b \vee d)\, P(ab \mid bd)$. An important special case of these formulas occurs when $ab \leq \neg d$ and $cd \leq \neg b$, like when b and d are disjoint. Then they both reduce to:

$$P((a|b) \vee (c|d)) = P((a|b) \wedge (c|d)) = P(b \mid b \vee d)\, P(a|b) + P(d \mid b \vee d)\, P(c|d). \tag{8}$$

More generally, if $b_1, b_2, b_3, \ldots, b_m$ are pairwise disjoint and denoting $(b_1 \vee b_2 \vee b_3 \vee \ldots \vee b_m)$ by b, then

$$P((a_1|b_1) \vee (a_2|b_2) \vee \ldots \vee (a_m|b_m)) = P(b_1|b)P(a_1|b_1) + P(b_2|b)P(a_2|b_2) + \ldots + P(b_m|b)P(a_m|b_m) \tag{9}$$

Applications

A probabilistically faithful algebra of conditional propositions will be useful for analyzing complex conditional propositions, for simplifying expert system rules, for analyzing iterated (nested) conditional constructions, for combining and simplifying data bases, and for analyzing common natural language constructions. In the following paragraphs somewhat more elaborate examples of these will be illustrated.

Complex Conditional Propositions

The disjunction or conjunction of conditionals naturally arises or is useful in various circumstances.

Communication Link Example. Suppose a military planner is estimating the conditional probability that both communication links, L_1 and L_2, that are connected in series, will survive if they are attacked. The planner therefore wants to calculate the probability that "if L_1 is attacked (b) or L_2 is attacked (d) then both L_1 will survive (a) and L_2 will survive (c)." This is just the conditional $(ac) \mid (b \vee d)$. Now

$$\begin{aligned}(ac) \mid (b \vee d) &= (ac) \mid (bd \vee \neg bd \vee b\neg d)\\ &= (ac \mid bd) \vee (ac \mid \neg bd) \vee (ac \mid b\neg d),\end{aligned}$$

and so by equation (9),

$$\begin{aligned}P((ac) \mid (b \vee d)) &= P(bd \mid b \vee d)\, P(ac \mid bd)\\ &+ P(\neg bd \mid (b \vee d))\, P(ac \mid \neg bd)\\ &+ P(b\neg d \mid (b \vee d))\, P(ac \mid b\neg d).\end{aligned}$$

Since $(a \mid \neg b) = (1 \mid \neg b)$, therefore $(a = 1) \mid \neg b$. So $(ac \mid \neg bd) = (ac \mid \neg b) \mid d = (c \mid \neg b) \mid d = (c \mid \neg bd) = (c|d) \mid (\neg b)$. If the survival of one link is (conditionally) independent of attack on, or survival of, the other link (a la J. Pearl [17]), then $P(\,(c|d) \mid (\neg b)\,) = P(c|d)$. So $P(ac \mid \neg bd) = P(c|d)$. Similarly $P(ac \mid b\neg d) = P(a|b)$ and $P(ac \mid bd) = P(a|b)P(c|d)$. So, since $P(a|\neg b) = 1 = P(c|\neg d)$,

$$\begin{aligned}P((ac) \mid (b \vee d)) &= P(bd \mid b \vee d)\, P(a|b)P(c|d)\\ &+ P(\neg bd \mid (b \vee d))\, P(c|d)\\ &+ P(b\neg d \mid (b \vee d))\, P(a|b),\end{aligned}$$

which is a formula whose values can be estimated according to the importance of each term.

If instead, the links are connected in parallel, then communication will survive as long as at least one link survives. So the relevant compound conditional is

$$\begin{aligned}(a \vee c) \mid (b \vee d) &= (a \vee c) \mid (bd) \vee (a \vee c) \mid (b\neg d \vee \neg bd)\\ &= (a \vee c) \mid (bd) \vee (1 \mid (b\neg d \vee \neg bd)).\end{aligned}$$

The second equality follows since one link is sure to survive if not both are attacked, and so by equation (9),

$$
\begin{aligned}
P((a \vee c) \mid (b \vee d)) &= P((a \vee c) \mid bd)\, P(bd \mid (b \vee d)) \\
&\qquad + (1)\, P((b\neg d \vee \neg bd) \mid (b \vee d)) \\
&= P((a \vee c) \mid bd)\, P(bd \mid (b \vee d)) + [1 - P(bd \mid (b \vee d))] \\
&= 1 - P(bd \mid (b \vee d))\, [1 - P((a \vee c) \mid bd)] \\
&= 1 - P(bd \mid (b \vee d))\, P(\neg a \neg c \mid bd),
\end{aligned}
$$

which is 1 minus the probability that both links will be attacked given one is, times the probability that both links will not suuvive given both are attacked. If link non-survivals are conditionally independent, then $P((a \vee c) \mid (b \vee d)) = 1 - P(bd \mid (b \vee d))\, P(\neg a \mid b) P(\neg c \mid d)$.

Expert System Example. Tin A. Nguyen et al [16] have given a nice account of the many difficulties arising in expert rule systems including circularity and consistency problems when a list of rules is left uncombined and unsimplified before chaining of the rules is attempted. Such a conjunction of "if - then - " rules can cause a computer to execute an infinite do-loop. For example, on page 72 they give an diagnostic example of three rules with a circularity problem: Denote the following propositions as indicated:

T = proposition that patient has a temperature > 100° F
F = proposition that patient has a fever
S = proposition that patient has flat pink spots
M = proposition that patient has measles.

Rule 1: If T	then F
Rule 2: If (F and S)	then M
Rule 3: If M	then T

Clearly, if these rules are chained, then rule 1 chains to rule 2, rule 2 chains to rule 3, and rule 3 chains back to rule 1, and the computer program may never end! But if these three rules are first conjoined by equation (5), $(F|T) \wedge (M \mid FS) \wedge (T \mid M) = (F|T) \wedge (T \mid M) \wedge (M \mid FS) = ((FT) \mid (T \vee M)) \wedge (M \mid FS) = FTM\neg(FS) \vee (\neg T \neg M)MFS \vee FTMS \mid (T \vee M \vee FS) = FT(\neg S \vee M) \mid (T \vee M \vee FS)$. So the three rules are equivalent to a single rule that "if a patient has a temperature over 100° F, or measles, or a fever and flat red spots, then the patient has a fever and a temperature over 100° F and either measles or no flat red spots." By thus combining the rules, there is no need to chain and no danger of an infinite loop. (Self-chaining can be disallowed.)

Throughout the above computation there is no assumption that the expert rules are wholly true; there may be a probability of each one holding, and these probabilities are not violated by the computations. Furthermore, the probability of any one, or of the conjunction of two, or of all three of the rules can be calculated in principle. For instance, the probability of all three is $P(FT\neg S \vee FTM) / P(T \vee M \vee FS)$.

Complex Conditional Example. Consider the familiar experiment of rolling a single die once and observing the value r facing up from 1 to 6. What is the conditional event (C|D) and conditional probability P(C|D) that [(if $3 \le r \le 5$ then r is odd) or (if $4 \le r$ then r is even)] and [if r is even then $r \le 4$]? In conditional notation, $[\,(\text{odd} \mid 3 \le r \le 5) \vee (\text{even} \mid 4 \le r)\,] \wedge [\,r \le 4 \mid \text{even}\,] = ?$ And what is its probability ?

Solution: First $[\,(\text{odd} \mid 3 \le r \le 5) \vee (\text{even} \mid 4 \le r)\,] = [(\text{odd})(3 \le r \le 5) \vee (\text{even})(4 \le r)] \mid [(3 \le r \le 5) \vee (4 \le r)] = [(3 \vee 5) \vee (4 \vee 6) \mid (3 \le r)] = [(3 \le r) \mid (3 \le r)]$. Furthermore, $[(3 \le r) \mid (3 \le r)] \wedge [(r \le 4) \mid \text{even}] = [(3 \le r) \vee (3 \le r)'] \wedge [(r \le 4) \vee (\text{even})'] \mid [(3 \le r) \vee \text{even}] = (r \le 5) \mid (2 \le r) = (2 \le r \le 5) \mid (2 \le r)$, which is (C|D) in reduced form. So P(C|D) = 4/5.

Example Distinguishing P(a) and P(a = 1). An experiment consists of flipping a fair coin once and then tossing either an ordinary 6-sided die once if "heads" comes up on the coin, or tossing a (4-sided) tetrahedron once if "tails" comes up on the coin. The faces of the die are numbered 1 through 6 and those of the tetrahedron 1 through 4. (The side of the tetrahedron touching the floor is chosen by the toss.) The set of possible outcomes is $\Omega = \{H1, H2, H3, H4, H5, H6, T1, T2, T3, T4\}$; $\mathfrak{B}$ is the collection of all subsets of Ω; and the probabilities assigned to Ω are $P(Hi) = (1/2)(1/6) = 1/12$ for i = 1 to 6, and $P(Ti) = (1/2)(1/4) = 1/8$ for i = 1 to 4. Let H denote {H1, H2, H3, H4, H5, H6} and let T denote {T1, T2, T3, T4}. The event (or proposition) "<5" is {H1, H2, H3, H4, T1, T2, T3, T4} and its probability is 1 - 2/12 = 5/6. Note that $P(<5 \mid H) = 2/3$ and $P(<5 \mid T) = 1$. Furthermore, P("<5" is certain | T) = P(("<5" = 1) | T) = 1 and P(("<5" = 1) | H) = 0. So P(<5 is certain after the coin-flip) = P("<5" = 1 after the coin flip) = P("<5" = 1) | H) P(H) + P("<5" = 1) | T) P(T) = 0 x P(H) + 1 x P(T) = P(T) = 1/2. Thus P("<5" after coin flip) = 5/6 but P("<5" is certain after coin flip) = 1/2.

Something very similar is going on in the example of E. Adams [1]. For instance, one can say above that the probability that "<5" is highly likely (in fact certain!) is 1/2 but the probability of "<5" is 5/6, which might not be considered "highly likely" and is anyway a very different probability.

Acknowledgement. I would sincerely like to thank Dr. Robert Manka of the National Research Council and Dr. Al Gordon and Ms. Sunny Conwell of NRaD for administrative support, and to thank all my colleagues doing research on this topic, especially I. R. Goodman, who has with unceasing energy, helped to develop this field and to provide me with critical analysis of my work.

References

[1] E. W. Adams, On the Logic of High Probability, *Journal of Philosophical Logic* **15**, 225-279, 1986.

[2] G. Boole, An Investigation of the Laws of Thought on which are Founded the Mathematical Theories of Logic and Probabilities, 2nd. ed., (Open Court, 1940) (1st. ed., Macmillan, 1854).

[3] P. G. Calabrese, The Probability that p Implies q (preliminary report), *Notices American Mathematical Society* 22 (3): A430-A431, April 1975.

[4] P. G. Calabrese, An Algebraic Synthesis of the Foundations of Logic and Probability, *Information Sciences* **42**, 187-237, 1987.

[5] P. G. Calabrese, Reasoning with Uncertainty Using Conditional Logic and Probability, in: *Proc. First International Symposium on Uncertainty Modeling and Analysis* by IEEE Computer Society, 682-688, 1990.

[6] P. G. Calabrese, Deduction and Inference Using Conditional Logic and Probability, Chapter 2 in *Conditional Logic in Expert Systems*, I. R. Goodman, et al, eds., North-Holland, 71-100, 1991.

[7] P. G. Calabrese, (Submitted for publication in *J. of Philosophical Logic*) Deductive Systems of Conditional Propositions in Conditional Probability Logic.

[8] B. De Finetti, *Probability, Induction and Statistics*, Wiley, New York, 1972.

[9] D. Dubois and H. Prade, Conditional Objects and Non-Monotonic Reasoning, *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, J. Allen, R. Fikes and E. Sandewall eds., Cambridge, Massachusetts, April 22-25, 1991.

[10] T. Hailperin, Probability Logic, *Notre Dame J. of Formal Logic* 25(3):198-212, 1984.

[11] J. N. Hooker, A Quantitative Approach to Logical Inference, Decision Support Systems (4) 45-69, 1988.

[12] I. R. Goodman, H. T. Nguyen and E. A. Walker, eds. *Conditional Information and Logic for Intelligent Systems: A Theory of Measure-Free Conditioning*, North-Holland, 1991.

[13] I. R. Goodman, H. T. Nguyen, M. M. Gupta, and G. S. Rogers, eds. *Conditional Logic in Expert Systems*, North-Holland, 1991.

[14] D. Lewis, Probabilities of conditionals and condi-tional probabilities, The Philos. Rev. **85** (3) 297-315, 1976.

[15] N. J. Nilsson, Probabilistic Logic, *Artificial Intelligence* **28**, 71-87, 1987.

[16] T. A. Nguyen, W. A. Perkins, T. J. Laffey and D. Pecora, Knowledge Base Verification, AI Magazine, 69-75, Summer 1987.

[17] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufman, San Mateo, 1988.

[18] B. Russell, and A. N. Whitehead, *Principia Mathematica*, Cambridge University Press, 1913.

[19] G. Schay, An Algebra of Conditional Events, *Journal of Mathematical Analysis and Applications* 24, 334-344, 1968.

[20] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[21] L. Zadeh, The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems, *J. Fuzzy Sets and Systems* (11), 199-227, 1984.

Philip G. Calabrese (born in Chicago on February 21, 1941) received the B.S., M.S. and Ph. D. degrees in mathematics (minoring in physics) in 1963, 1965 and 1968 respectively from Illinois Institute of Technology in Chicago. His masters work in boolean logic was published in the *Notre Dame J. of Formal Logic* and his Ph. D. work in probabilistic geometry was published in *Aequationes Mathematicae*.

In 1968 he became a faculty member in the Department of Mathematics at the U. S. Naval Postgraduate School in Monterey, California with teaching and research responsibilities. He then became a founding faculty member at the new experimental California State College at Bakersfield from 1970 to 1973, and also during 1975, designing and teaching undergraduate mathematics courses. During the 1974 academic year, as a Visiting Scholar to the University of California at Santa Barbara, he was able to develop some of his early ideas for combining logic and probability and to explore other fundamental concepts. In 1975 he became a faculty member in the Department of Mathematics at Humboldt State University in Arcata, California and later became Director of the Alternative Energy Project in the School of Interdisciplinary Studies, where his students made energy modifications on campus buildings. During this period he initiated his engineering consulting firm, Data Synthesis, and was certified by the State of California as an Energy Analyst to do energy calculations for commercial and residential structures. In 1979 he took a position in the aerospace industry as an analyst at Decision Science, Inc., in San Diego where he performed large scale, multi-attribute utility analysis and mathematical modeling, including FORTRAN computer programming. In 1982, he accepted a position as principle investigator at Orincon Corp. in San Diego to do acoustic signal processing research. He moved to Logicon Corp. in 1983, where he helped design and implement large scale, high fidelity computer simulations of radar surveillance systems for the U. S. Department of Defense. In 1987, he published in *Information Science* a long foundations paper "An Algebraic Synthesis of the Foundations of Logic and Probability". In 1990, shortly after leaving Logicon, he won a full-time award as a Senior Research Associate of the National Research Council to do research in his specialty of conditional probability logic at the Research, Development, Test & Evaluation Division (NRaD) of the Naval Command, Control and Ocean Surveillance Center in San Diego.

Dr. Calabrese early showed interest in mathematical research, publishing in the *American Mathematical Monthly* an original paper "A note on alternating series" as a college sophomore in calculus. This paper was subsequently anthologized in *Selected Papers in Calculus* by T. Apostol et al. He was offered full-time fellowship for graduate study in mathematics at the University of California at Berkeley in 1963, but instead accepted a 3-year NASA Traineeship at Illinois Institute of Technology. He has appeared in Who's Who in America and American Men and Women of Science. He was inducted into the Sigma Xi and Sigma Pi Sigma science and physics honor societies. Dr. Calabrese and his wife Cindy of 20 years are raising four children.